Everything I Disagree With is #FakeNews:

Correlating Political Polarization and Spread of Misinformation

Manoel Horta Ribeiro
Pedro H. Calais
Virgílio A. F. Almeida
Wagner Meira Jr.

UNIVERSIDADE FEDERAL DE MINAS GERAIS

DCC
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

- News consumption after Online Social Networks:

Reputation
matters less

Profits comes
from clicks

Recommended
content

- Due to this, two phenomena have their impact increased:

Opinion Polarization

Spread of Misinformation

Opinion Polarization

"the extent to which opinions on an issue are opposed in relation to some theoretical maximum"

- Recommendation algorithms may limit users to ideologically diverse content.
- System may fuel partisan news, thus increasing polarization.

Motivation

"misinformation is false or incorrect information, that is spread intentionally or unintentionally"

Spread of Misinformation

- Made easier by the decrease in the accountability of sources.
- Bots may be employed to disseminate misinformation.

- The media suggests an interaction between these:

Opinion
Polarization

Spread of
Misinformation

- The media suggests an int

P

Stop Calling Everything “Fake News”

Journalists are blurring several problems into one—and making it impossible to solve.

By Will Oremus

3k 574 619

Spread of Misinformation

- But also previous studies also do:

Opinion Polarization

polarized groups are more susceptible to the dissemination of misinformation

Spread of Misinformation

- But also previous studies also do:

Opinion Polarization

the dissemination of misinformation plays a key role in creating polarized groups

Spread of Misinformation

- Can this interaction happen in some other way?

Motivation

"everything I disagree with is fake news"

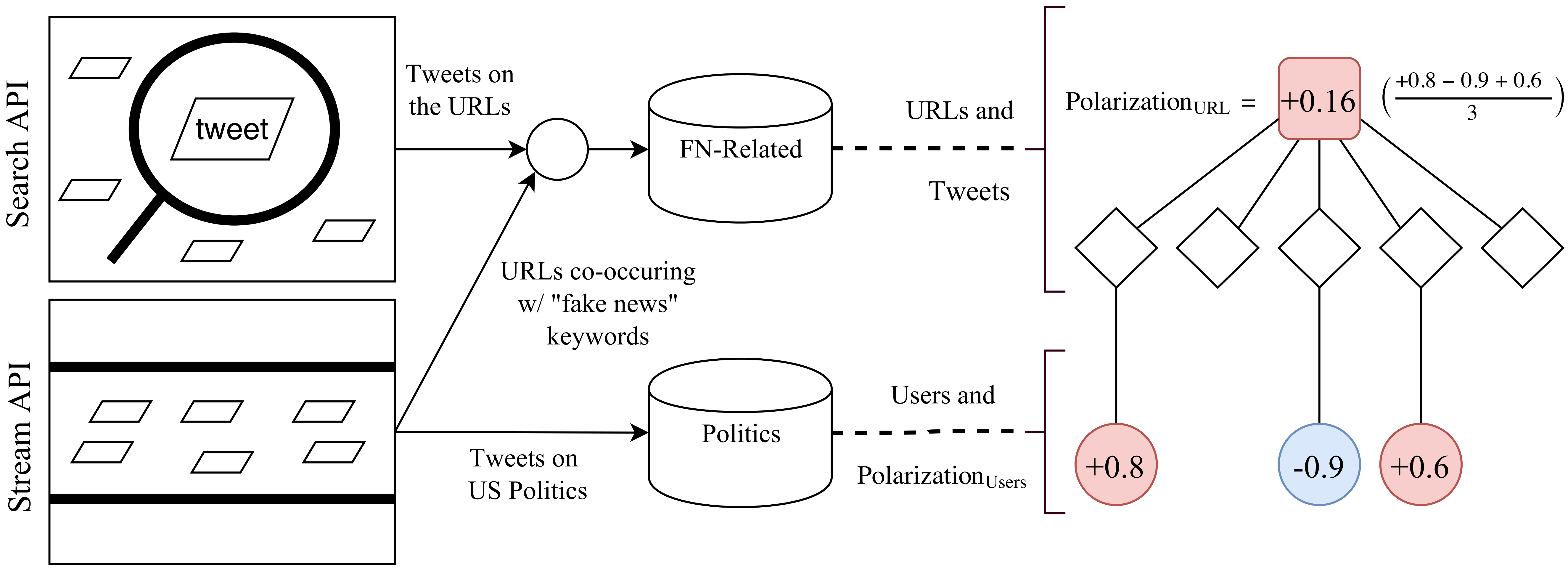
???

- Users designate incorrectly classify sources of misinformation due to disagreement.
- Alternate narratives of "what is true"

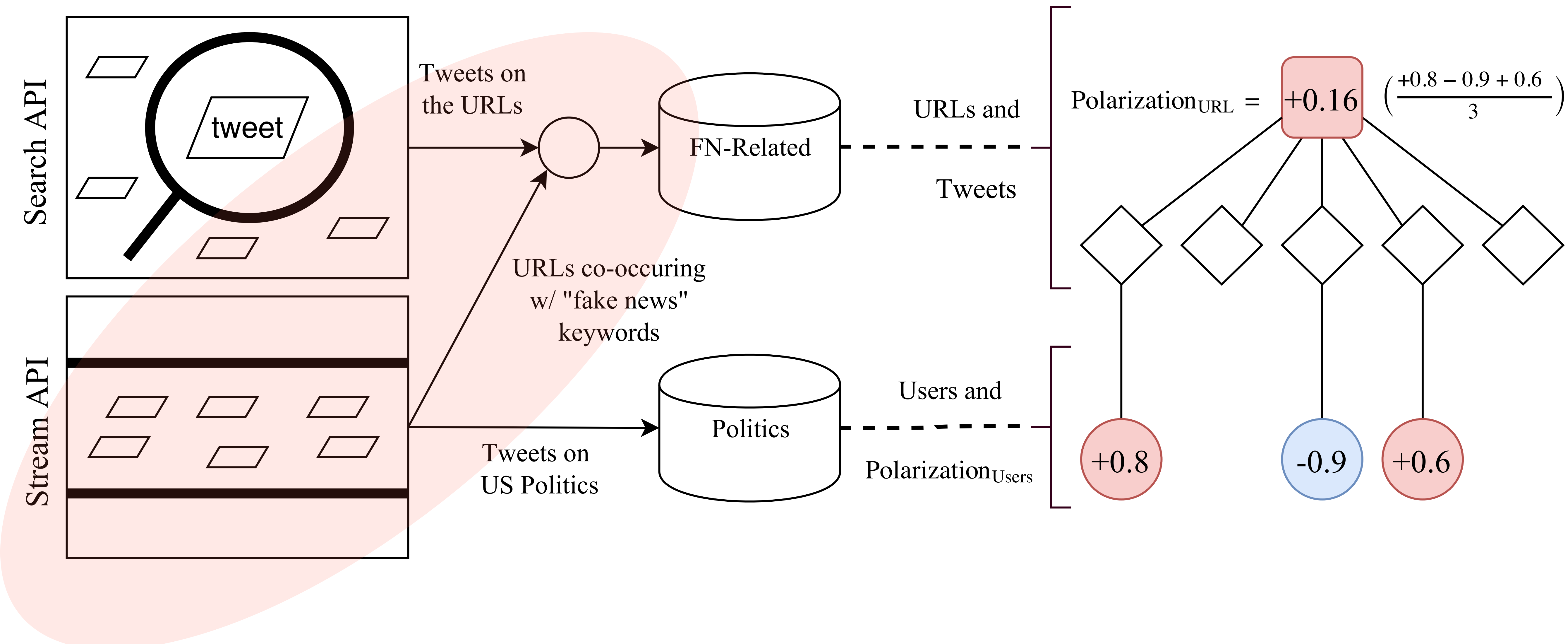
Q1: How is polarization quantitatively related to information perceived as or related to fake news?

***Q2*:** Are users designating content that they disagree with as misinformation?

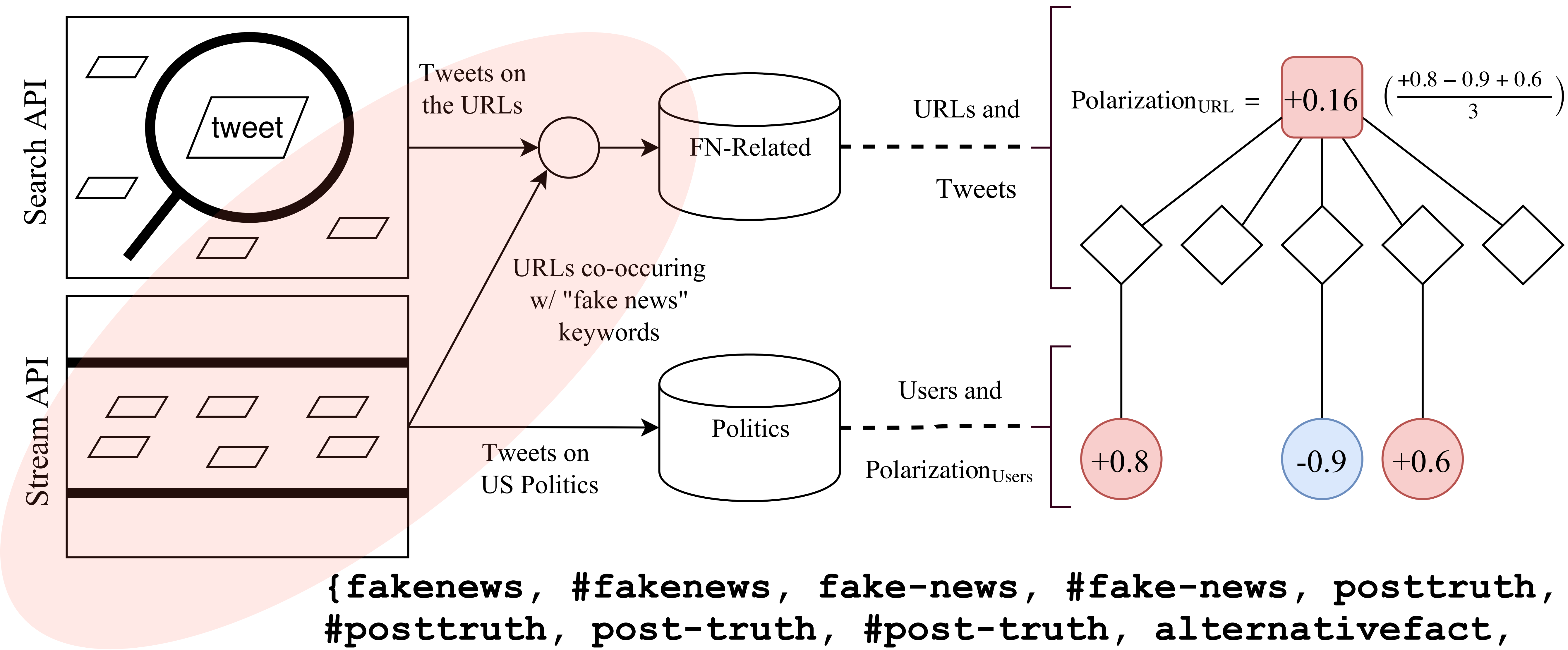
- We collect a dataset trying to answer this questions in the following fashion:



- We collect the *tweets* with words and hashtags related to misinformation using the stream API.



- We collect the *tweets* with words and hashtags related to misinformation using the stream API.

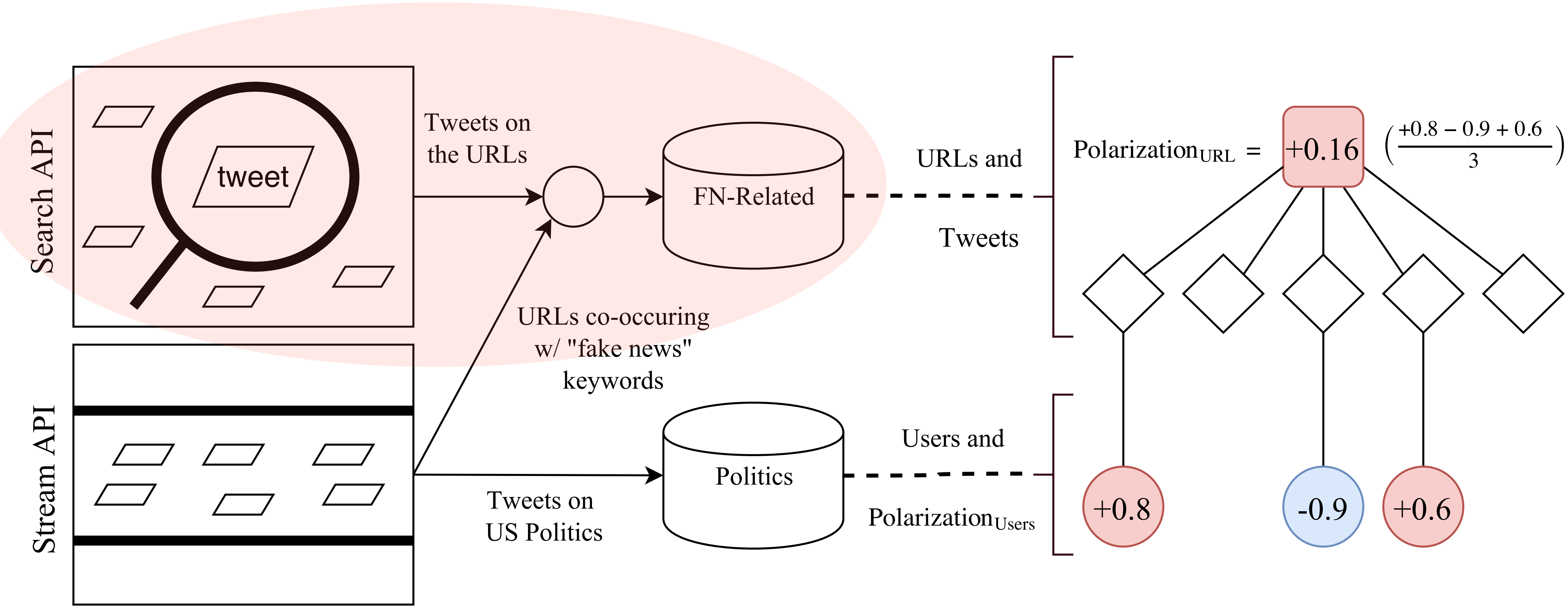


{fakenews, #fakenews, fake-news, #fake-news, posttruth, #posttruth, post-truth, #post-truth, alternativefact, #alternativefact, alternative-fact, #alternative-fact}

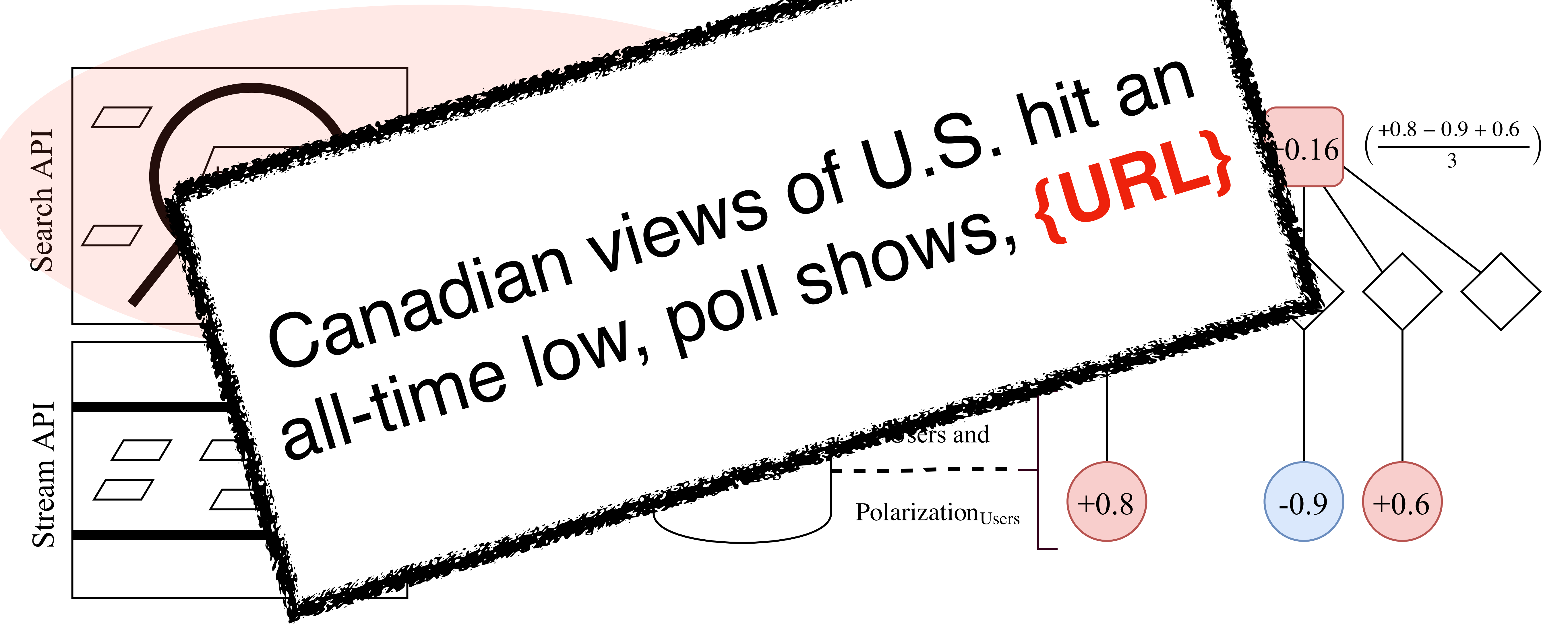
- We collect the *tweets* with words and hashta... misinformation using the stream API...

{fakenews, #fakenews, fake-news, #fake-news, posttruth, #posttruth, post-truth, #post-truth, alternativefact, #alternativefact, alternative-fact, #alternative-fact}

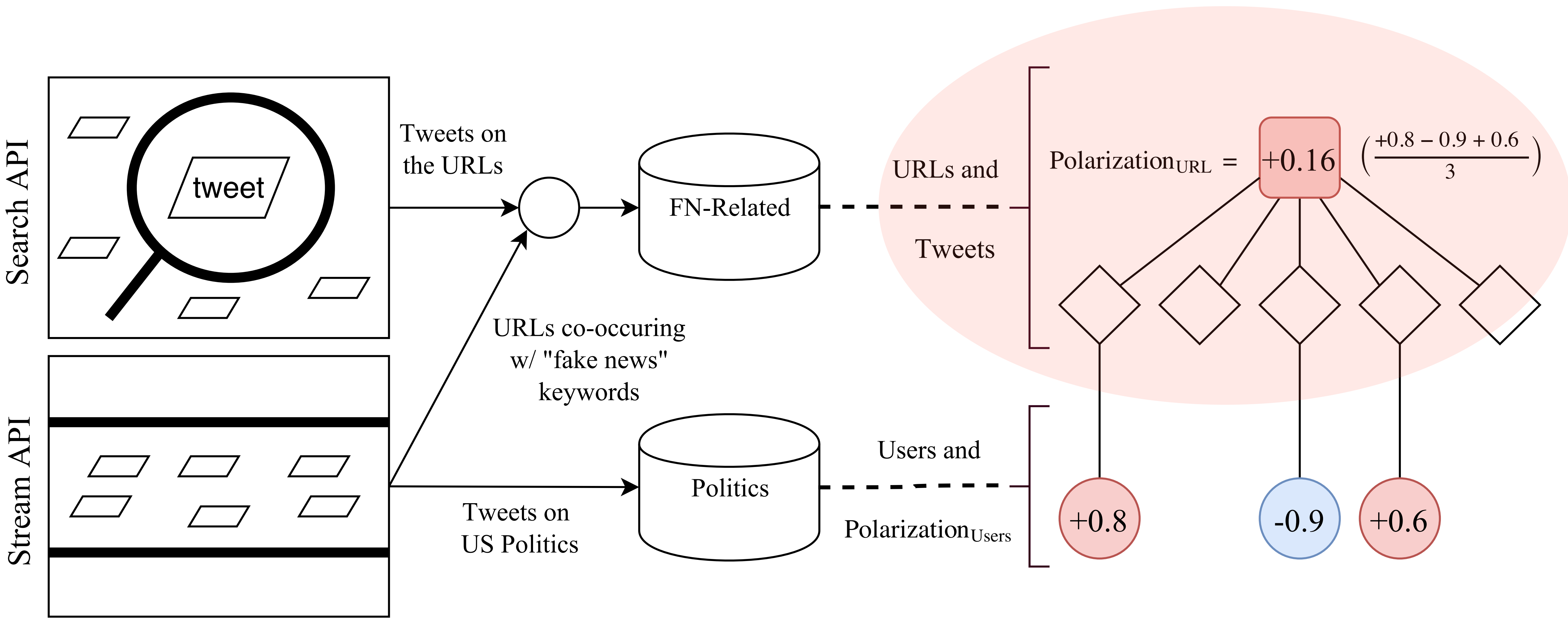
- We use the URLs in the tweets in the search API and find more general tweets about it (not necess. w/ keywords)



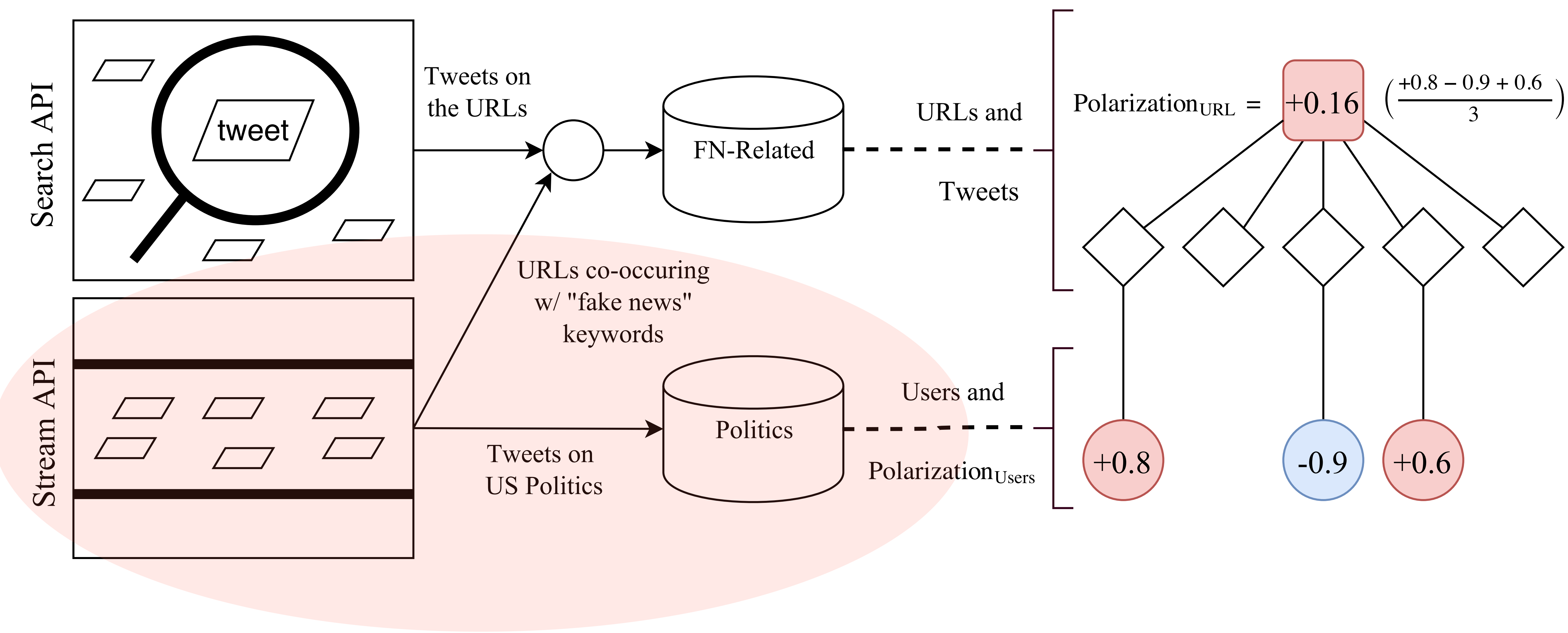
- We use the URLs in the tweets in the search API and find more general tweets about it (not ne ywords)



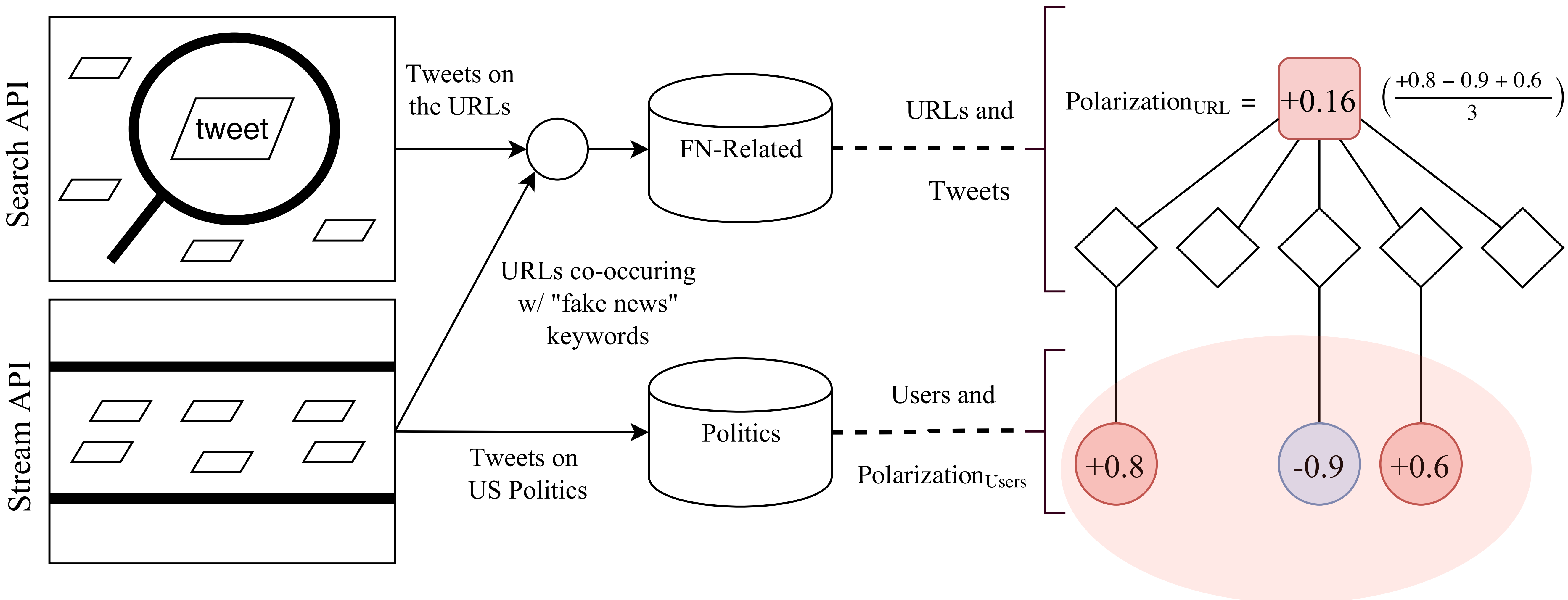
- With this we can manage to get an URL and a many associated tweets.

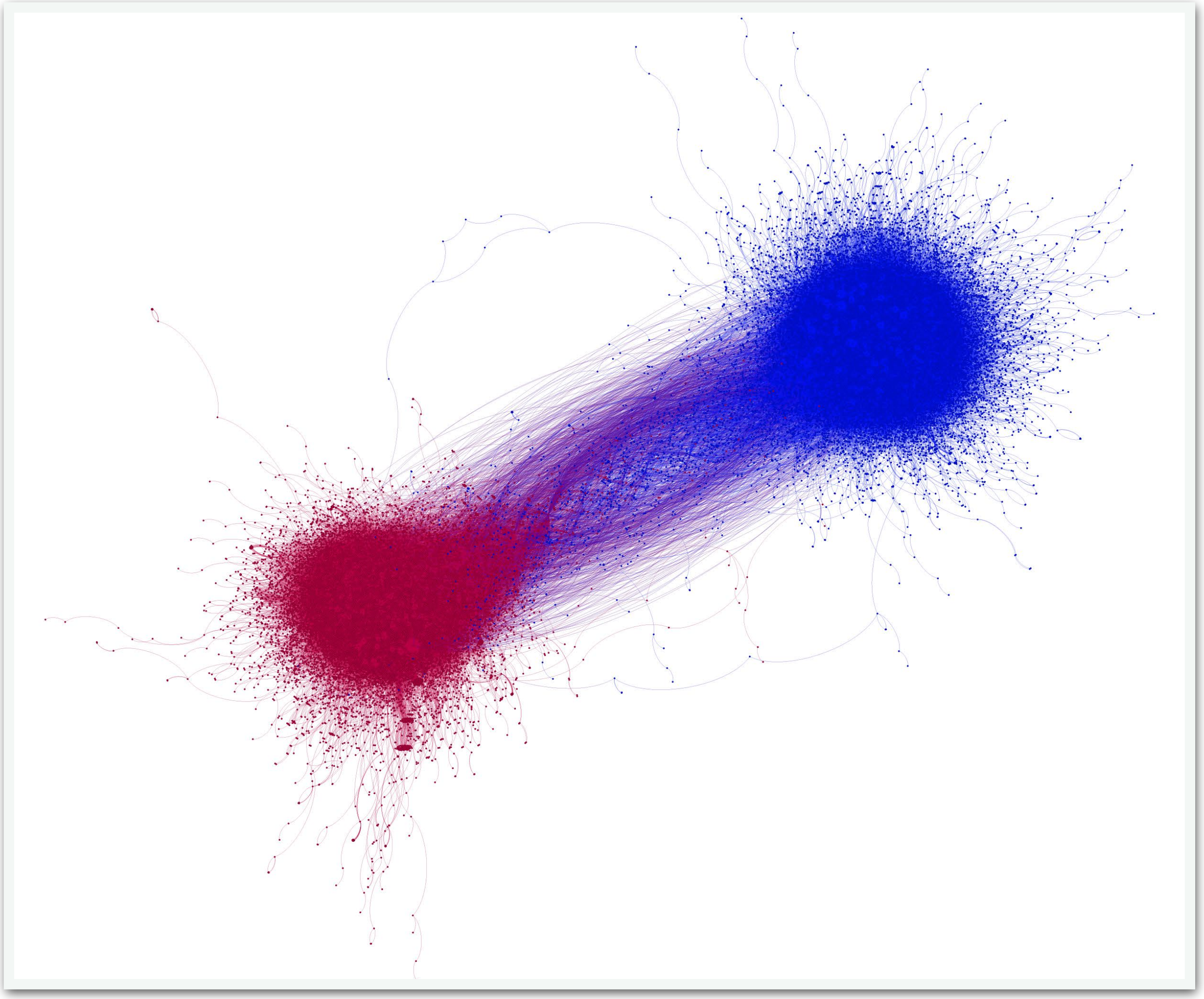


- The second step envolves a bigger data collection in the stream API involving more broad political hashtags



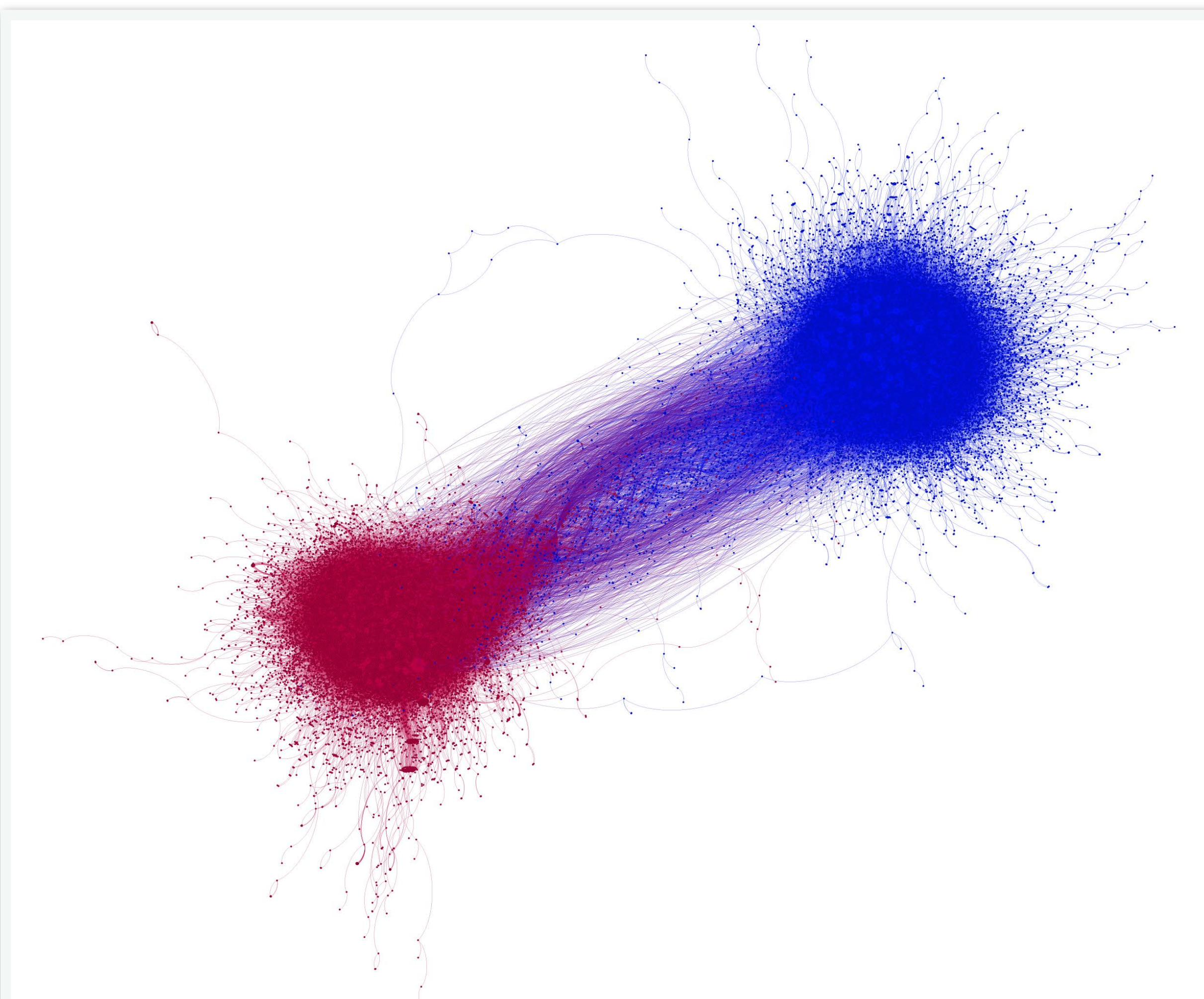
- This allow us to (with an community detection algorithm) get a polarization metric for some of the users





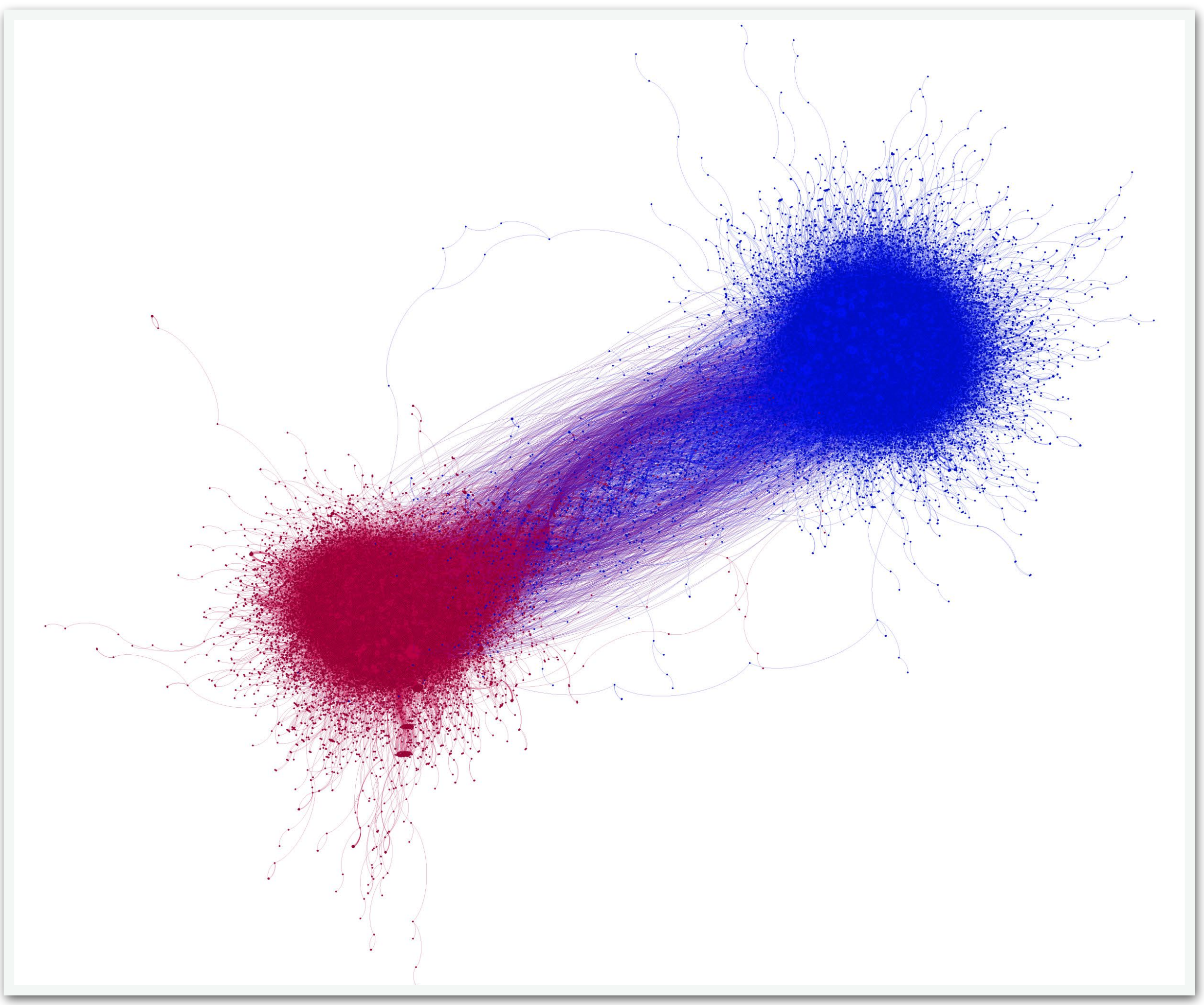
- Assume that the number of communities K formed around a topic T is known

- We build the retweet bipartite graph using the retweets in the collected dataset.



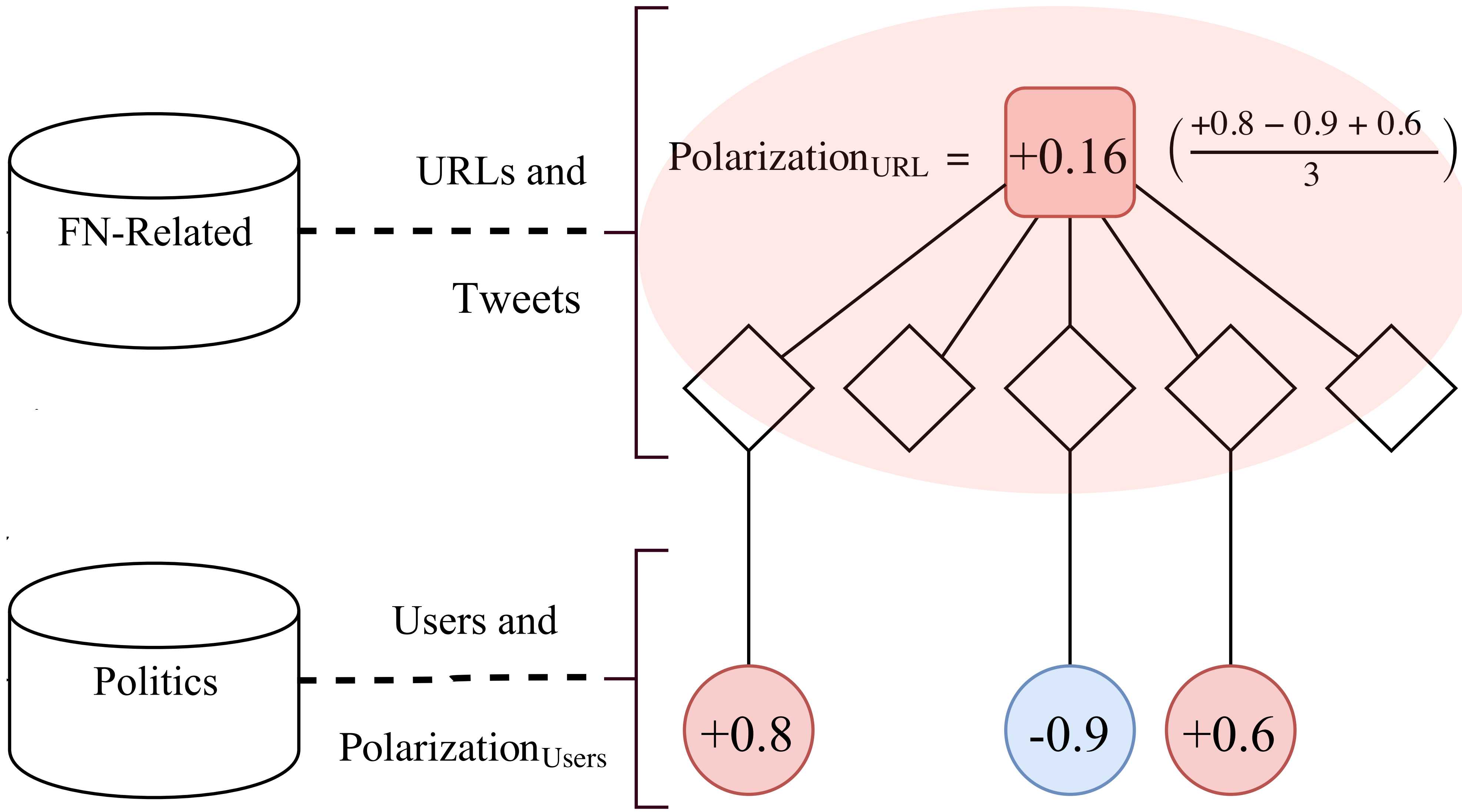
- We select seeds with known political position, (i.e. politicians)

- A random walker departs from each seed and travels, with some probability of restarting from its original

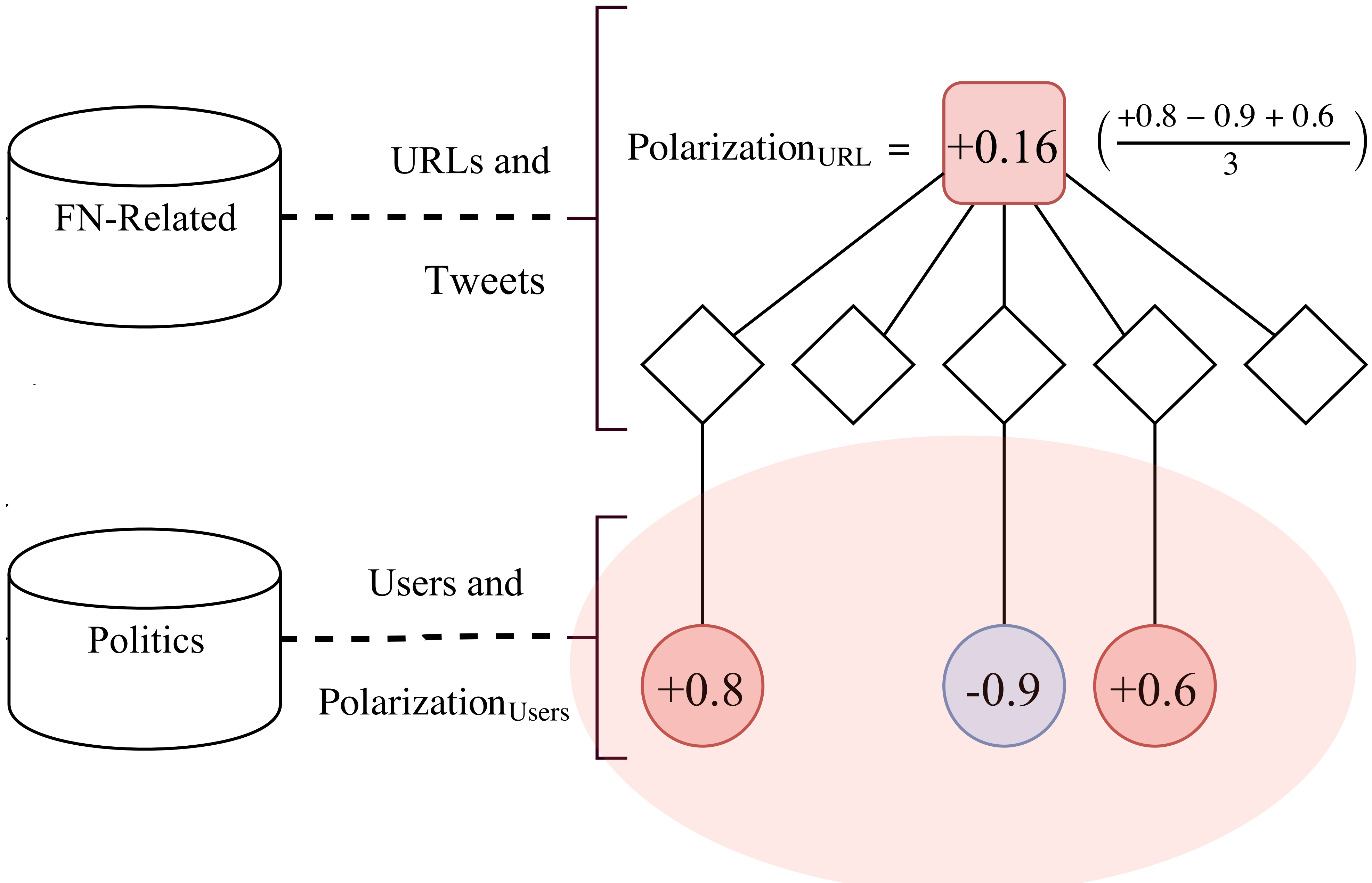


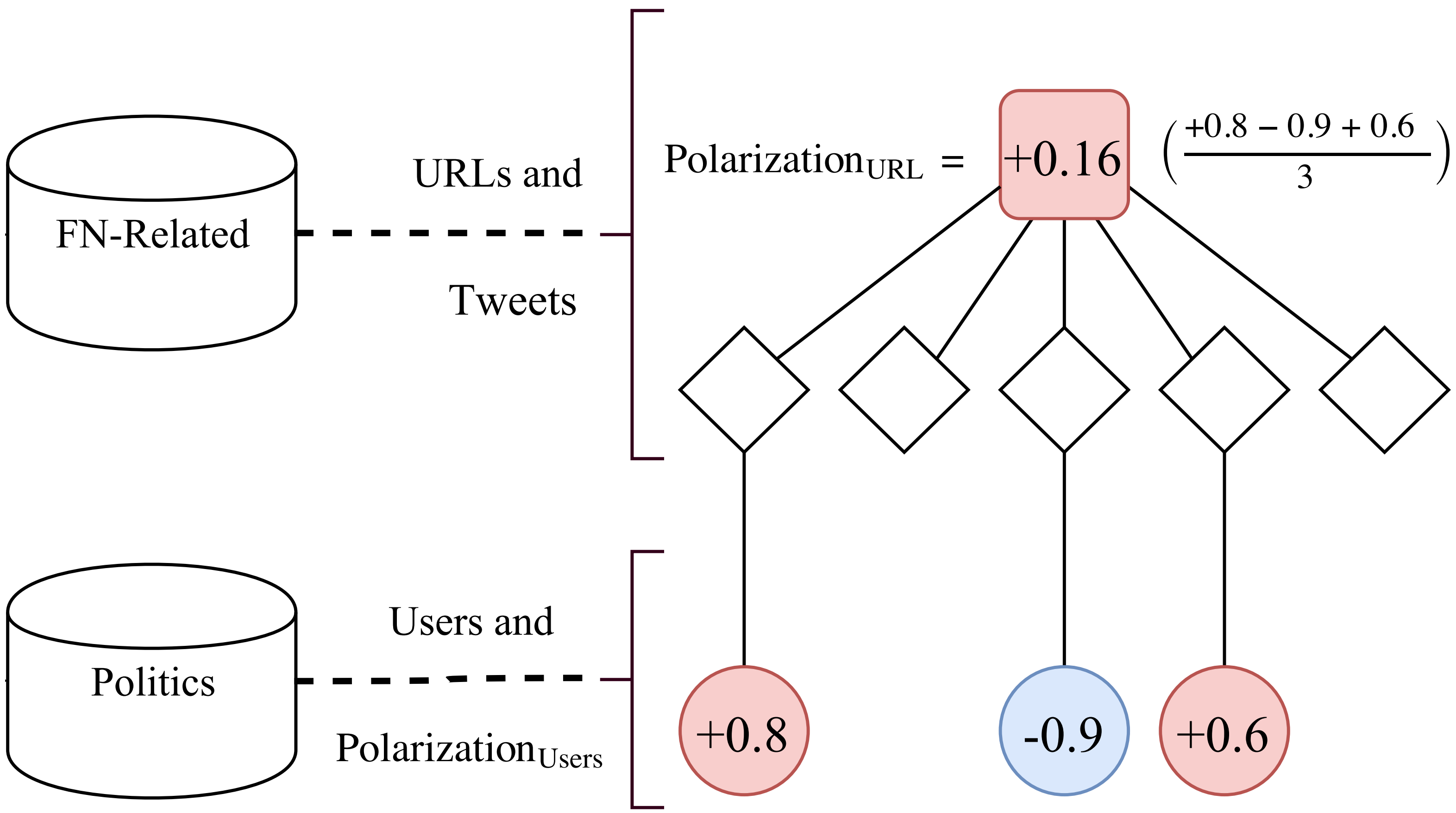
- The relative proximity of each node to the sets of seeds yield a prob. that that node belongs to that community

Method



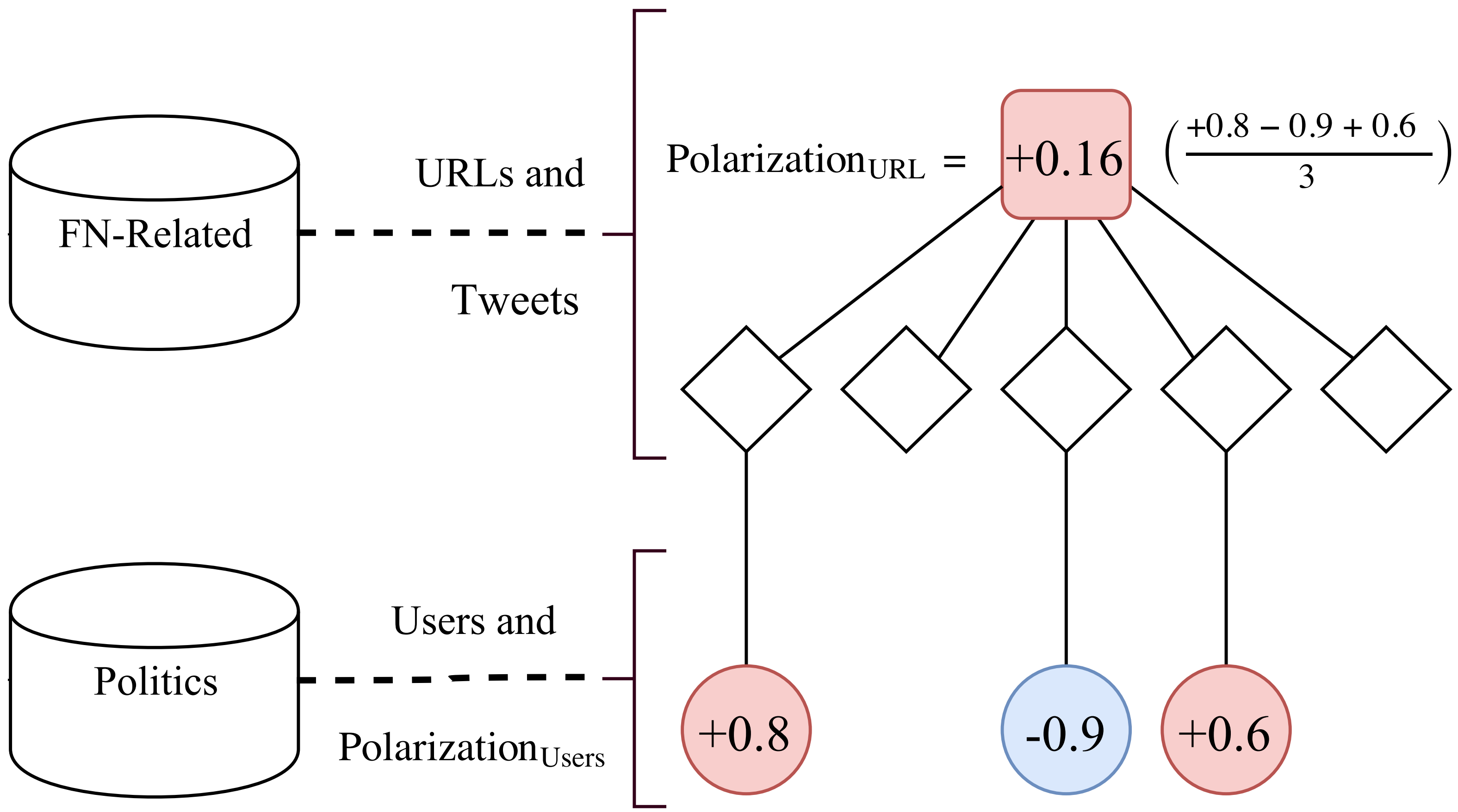
FN-Related
URLs and
Tweets
Politics
Users and
Polarization$_{Users}$
Polarization$_{URL}$ = +0.16 $\left(\frac{+0.8 - 0.9 + 0.6}{3}\right)$
+0.8
-0.9
+0.6


FN-Related
URLs and
Tweets
Politics
Users and
PolarizationUsers
PolarizationURL =
+0.16
(+0.8 − 0.9 + 0.6 / 3)
+0.8
-0.9
+0.6



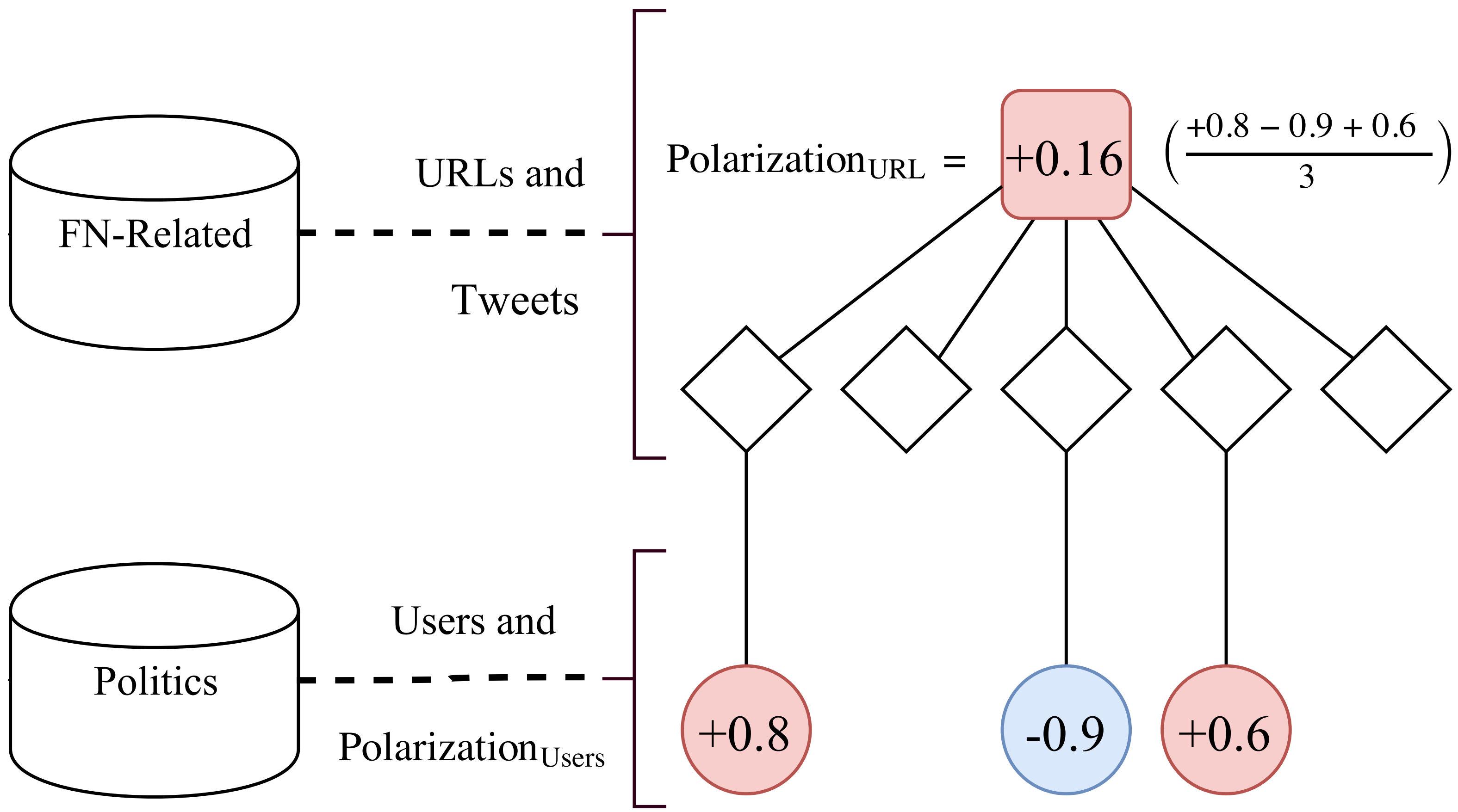
- With this data we:

(i) Estimate users political polarization on different domains.



- With this data we:

(ii) Estimate political polarization of URLs.



- With this data we:

(iii) Qualitatively analyze the domains and the content of the URLs.

- Data collection extracts the political orientation of 374,191 of users that commented some of the collected URL (29%)

	General Statistics				Shared Users		Shared Active Users	
Source	#users	#active users	#tweets	#urls	FN-Related	Politics	FN-Related	FN-Related
FN-Related	374,191	101,031	833,962	109,397	-	29.22%	-	37, 61%
Politics	4,164,604	247,435	246,103,385	-	2.62%	-	15.72%	-

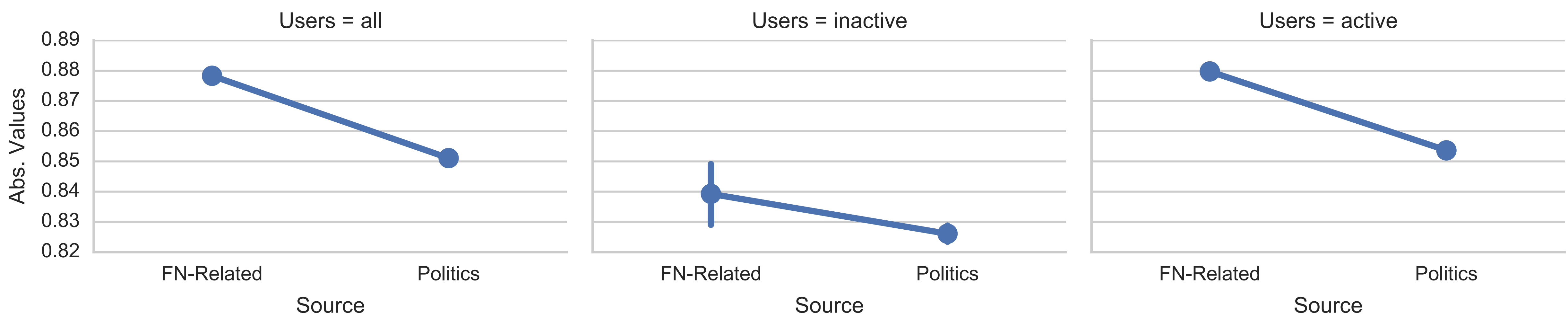
Table 1: General characterization of the data sources. The intersection between the Politics dataset and FN-Related is important as we use it to characterize the polarization of the users, and consequently of the URLs in the FN-Related datasets.

- Although it is a relatively small sample of all the users in a broader context (2.67%), it jumps to 15.72% when we consider only the active users.

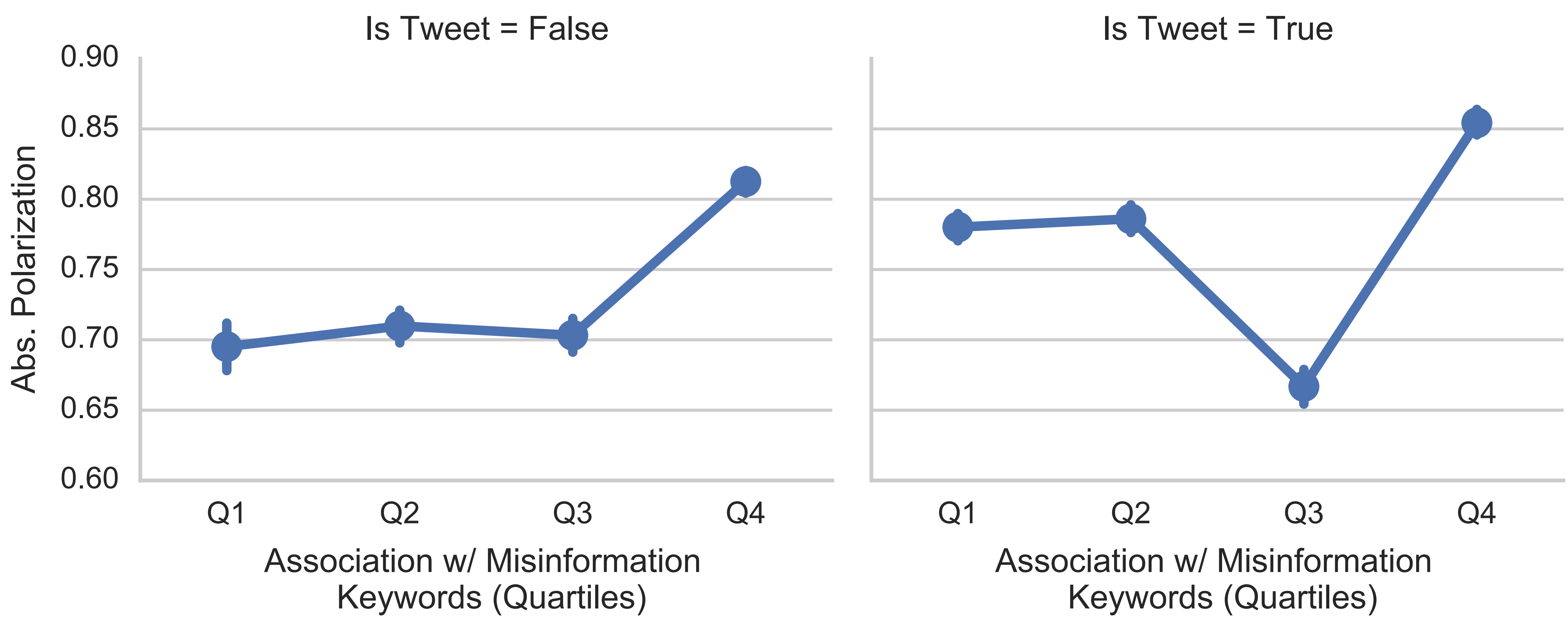
	General Statistics				Shared Users		Shared Active Users	
Source	#users	#active users	#tweets	#urls	FN-Related	Politics	FN-Related	FN-Related
FN-Related	374,191	101,031	833,962	109,397	-	29.22%	-	37, 61%
Politics	4,164,604	247,435	246,103,385	-	2.62%	-	15.72%	-

Table 1: General characterization of the data sources. The intersection between the `Politics` dataset and `FN-Related` is important as we use it to characterize the polarization of the users, and consequently of the URLs in the `FN-Related` datasets.

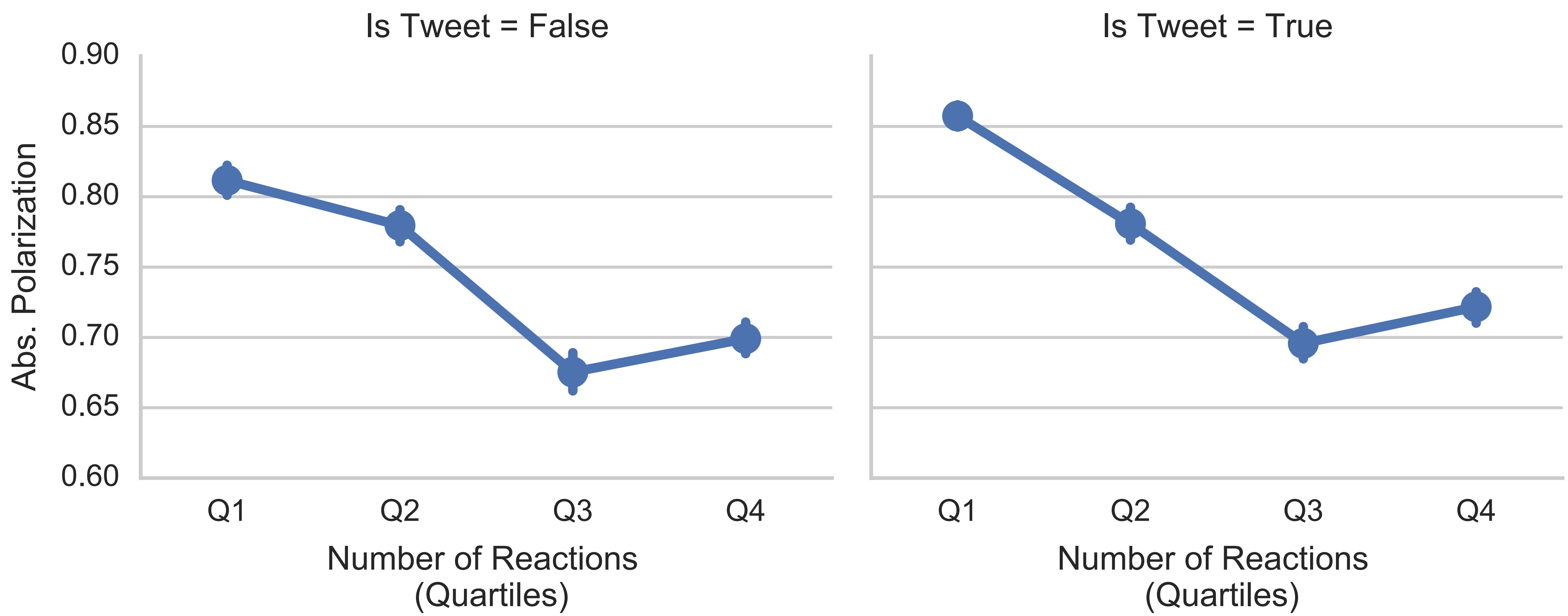
- The users in the fake-news-related dataset are more polarized than in the general politics one. This is evidence that fake-news-related discourse induces polarization.



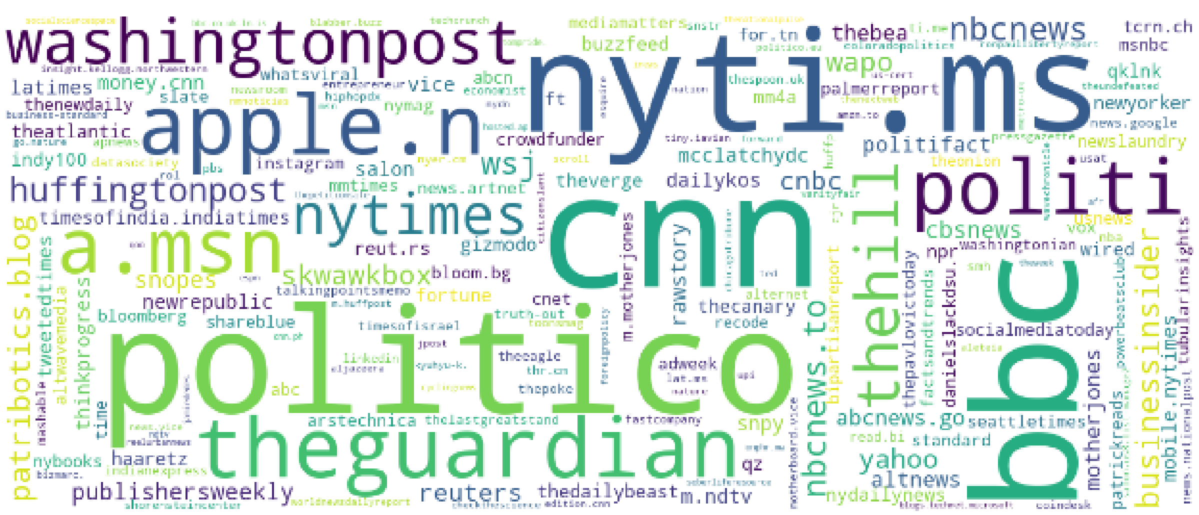
- The polarization grows according to association with misinformation.

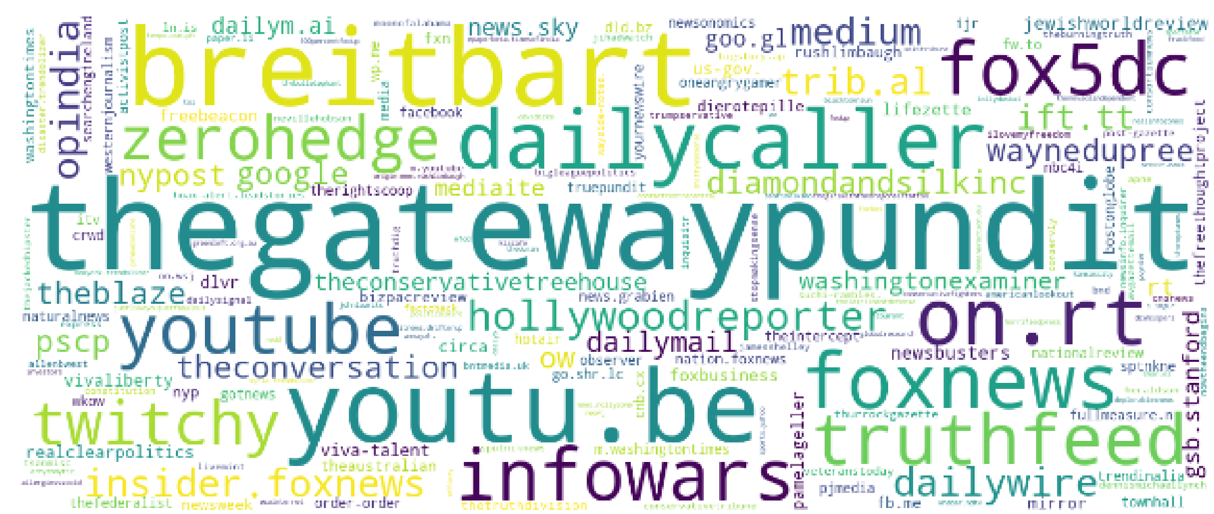


- The polarization decreases with number of reactions.



- People cite sources that they agree ideologically with in this fake-news-related context.





- Qualitatively analyzing top URLs.

{DISMISSING A NARRATIVE}
New York Post: FBI clears
Michael Flynn in probe
linking him to Russia

- Qualitatively analyzing top URLs.

{HUMOUR}
Telegraph: Prisoner dressed as woman in failed escape bid

- Qualitatively analyzing top URLs.

{NEWS FAKE TAGGING}
Raw Story: Family blasts right-wing media for spreading fake news story about slain DNC staffer as Russia scandal deepens

- Qualitatively analyzing top URLs.

{URL FAKE TAGGING}
@realdonaldtrump: The Russia-Trump collusion story is a total hoax, when will this taxpayer funded charade end?

- We present quantitative evidence of various interactions of polarization and misinformation.
- We present qualitative evidence of different uses of misinformation-related tags.

What does this means?

Discussion

- This may present challenges for solutions that use the "wisdom of the crowd" to determine what is fake.
- Polarization may prove itself useful as a feature to distinguish between fake and biased.

- **Future directions**:
  - How to quantify the influence of bias on what is perceived as fake?
  - How to explicitly tell how biased a piece of information is? Should we do this?

Thank You!

github manoelhortaribeiro
twitter manoelribeiro
mail manoelribeiro at dcc.ufmg.br

DCC
DEPARTAMENTO DE
CIÊNCIA DA COMPUTAÇÃO

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INCIPIT VITA NOVA
7 DE SETEMBRO DE 1927